

EM 432 569514 U.S.

"Express Mail" mailing label number

3/19/98

Date of Deposit

I hereby certify that this correspondence is being deposited with the United States Postal Service "Express Mail Post Office to Addressee" service under 37 C.F.R. § 1.10 on the date indicated above and is addressed to the Assistant Commissioner for Patents, Washington, D.C. 20231

David Corley

(Typed or printed name of person mailing paper or fee)

David Corley

Signature of person mailing paper or fee)

Our Case No. 8998/3

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE
APPLICATION FOR UNITED STATES LETTERS PATENT

INVENTOR:

James A. Radosevich, Ph.D.

TITLE:

A GENE ENCODING A NOVEL
MARKER FOR CANCER

ATTORNEY:

Alice O. Martin
BRINKS HOFER GILSON & LIONE
P.O. BOX 10395
CHICAGO, ILLINOIS 60610
(312) 321-4200

60610-10395

F Cancer Marker Protein and Peptides thereof

~~A GENE ENCODING A NOVEL MARKER FOR CANCER~~

Inventor: James A. Radosevich, Ph.D.

5 The invention relates to a gene encoding a protein that is a cancer associated antigen useful as a marker that is not restricted to previously defined histological classes of cancer. Peptides of the antigen are useful as a vaccine for treatment and prevention of cancer, and for the preparation of new, specific, monoclonal antibodies.

10 BACKGROUND OF THE INVENTION

Cancer¹ is a leading cause of death in men and women throughout the world. In the United States alone, over 1 million new cases are diagnosed each year, and over 0.5 million deaths are reported annually (Landis, *et al.*, 1998). Historically, tumors are grouped and treated, based in part by the tissues in which they arise, e.g.: breast cancer, colon cancer, and lung cancer, and the like. Yet, within lung cancer, for example, it is well recognized that these tumors are a very heterogeneous group of neoplasms. This is also true for tumors arising in other tissues. In part, because of this heterogeneity, there are complex and inconsistent classification schemes which are used for human tumors. Previous attempts to treat cancer have been hampered by : 1) the arbitrary classification of tumors arising within given tissues, and 2) by using microscopic methods based on how these tumors look (histological classification). Although existing classifications for

¹ Terminology used herein is as follows: "cancer" is a malignant tumor, wherein a "tumor" is an abnormal mass of tissue, that need not be malignant. "neoplasm" is a form of new growth.

various tumor types have some prognostic value, almost all of the classifications fail to predict responsiveness to treatments and likelihood of cure or disease course. Improved classification schemes based on the biological constitution of these neoplasms is required to significantly alter the survival statistics of humans who have cancer. One approach to solving these problems is to locate molecules specific to tumors, preferably antigens in molecules that are markers for cancer cells. (A "marker" is defined herein as any property which can be used to distinguish cancer from normal tissues and from other disease states.) The markers' presence is then a basis for classification.

Monoclonal antibodies (MCAs) prepared by somatic cell hybridization techniques, usually in mice, are useful molecular probes for the detection and discrimination of cellular antigens, and therefore have great potential for detecting cancer associated antigens. These antibodies bind to specific antigens and the binding is detectable by well known methods. When binding occurs, the inference is made that a specific antigen is present. Those cancer associated antigens which are exposed to the cell surface or found in the cancer mass, are molecular targets for the immune systems (including host antibodies) of the host. Recent findings suggest that cancer patients who have antibodies against their tumors, do better than those who do not mount this type of immune response (Livingston, *et al.*, 1994). Therefore, natural, induced, or administered antibodies are a promising therapeutic approach.

The humanization of non-human MCAs (the process by which non-human MCA reactive sites are shuttled into cloned human antibodies and expressed) results in reduced immunogenicity of the foreign antibodies without the loss of their specific binding in *in vivo* and *in ex vivo* applications. MCAs can be used as *in vivo* imaging agents, diagnostic tests, and for therapy (Radosevich, *et al.* 1988, 1990; Rosen, *et al.* 1988).

Vaccine therapy is a well established approach directed at inducing an immune response without exposure to the causative agent of a disease or

condition. Many vaccines are available, for example, to stimulate a response in a host to bacterial and viral agents. The use of tumor associated antigens (markers) in a vaccine could prevent primary cancer occurrence, and could also provide a means to prevent recurrence of the disease.

5 Gene therapy is a means by which the genetic make-up of cells is modified to express the gene of interest. There are many forms of gene therapy including: gene replacement, antisense suppression therapy, and surrogate gene expression. Discovering genes encoding cancer-associated, preferably cancer-specific antigens (markers) opens the door to genetic
10 intervention against cancer cell proliferation. The accurate and consistent use of a cancer marker to differentiate cancerous from normal tissue, not only has diagnostic potential, but is also desirable for treatment and prognosis. Therefore, such markers have been sought.

15 Recent studies have shown that the enzyme encoding human aspartyl beta-hydroxylase (HAAH) is overexpressed in some human adenocarcinoma cell lines, and in primary hepatocellular cancers, therefore could be a marker. The gene said to encode HAAH has been cloned and sequenced (Gronke, *et al.*, 1989, 1990; Wang, *et al.*, 1991; Jia, *et al.*, 1992, 1994; Koriath, *et al.*, 1994; Lavaissiere, *et al.*, 1996). However, little is known about HAAH
20 expression in human tumors in general (Lavaissiere, *et al.*, 1996).

25 The study of the HAAH enzyme grew out of the study of its bovine counterpart (Gronke, *et al.*, 1989, 1990; Wang, *et al.*, 1991; Jia, *et al.*, 1992). Bovine aspartyl beta-hydroxylase is an intracellular, glycosylated protein, localized in the rough endoplasmic reticulum. The protein has been reported
30 to have three major species of molecules; a 85 kilodalton form, and two active forms with molecular weights of 56 and 52 kilodaltons respectively (Lavaissiere, *et al.*, 1996).

 Using standard biochemical methods, bovine aspartyl beta-hydroxylase (bAAH) has been purified and characterized (Gronke, *et al.* 1990; Wang, *et al.*, 1991). The activity of the enzyme has been shown to be correlated with the 52 and 56 kilodalton species which were purified.

Immunologically, a related higher molecular weight form (85-90 kilodalton) was also observed. As part of the purification, bAAH is bound to Con A sepharose, which is consistent with the conclusion that the enzyme is glycosylated. (Subsequent reports on the DNA sequence show three possible glycosylation sites, with one site being very close to the known active enzyme domain.) The protein is very acidic in nature, and a detergent is not required to solubilize the active fraction. The active enzyme site is dependent from the biochemically isolated bovine protein (bAAH) on the presence of histidine at position 675 (Jia, *et al.*, 1994).

A partial amino acid sequence was obtained for HAAH. DNA probes (a DNA probe is a molecule having a nucleotide sequence that is capable of binding to a specified nucleotide sequence) deduced from this amino acid sequence was used to screen a bovine cDNA library (Jia, *et al.*, 1992). (A cDNA library contains the sections of DNA that encode for gene products, e.g. peptides, as opposed to genomic DNA). Several overlapping cDNA sequences in the library contained a 754 amino acid open reading frame (ORF) sequence which would be expected to encode an 85 kilodalton protein. Also present in this ORF sequence were two other possible start codons, that is, locations at which encoding begins. The most 3' start codon was preceded by a ribosome binding site. Translation of the clone having this sequence resulted in a protein that was about 85 kilodaltons. Antiserum was raised to the membrane fraction of human MG-63 cells and was used to immunoscreen a cDNA library made from MG-63 cells. Data on one clone was reported which could encode a 757 amino acid protein, and, by sequence analysis, was found to have strong N-terminal homology with bAAH (Korioth, *et al.*, 1994). When this clone was used in an *in vitro* translation system (an artificial cocktail of normal cell cytoplasm used to convert mRNA into protein), a 56 kilodalton protein was produced. It was suggested that this was due to posttranslational cleavage.

The HAAH enzyme is responsible for the modification of specific aspartic acid residues within the epidermal growth factor-like domains of

proteins. It has been hypothesized that these modified aspartic acid residues allow the epidermal growth factor-like domains to become calcium binding domains. (Gronke, *et al.*, 1989, 1990; Wang, *et al.*, 1991; Jia, *et al.*, 1992, 1994; Koriath, *et al.*, 1994; Lavaissiere, *et al.*, 1996).

5 An enzyme related to HAAH, aspartyl beta-hydroxylase (AAH), was first studied because it specifically modified select aspartic acid or asparagine residues in a group of biologically important proteins including the vitamin K-dependent coagulation factors VII, IX, and X. Other proteins like C, S, and Z also have this modification (Gronke, *et al.*, 1989, 1990; Wang, *et al.*, 1991; Jia, *et al.*, 1992, 1994; Koriath, *et al.*, 1994; Lavaissiere, *et al.*, 1996). Aspartic acid and asparagine residues have been shown to be modified by HAAH in proteins containing epidermal growth factor-like domains. The function of the beta-hydroxyaspartic and beta-hydroxyasparagine residues is unknown, however, it has been speculated that this modification is required for calcium binding in the epidermal growth factor EGF-like domains of selected proteins.

Antibodies were raised to human hepatocellular carcinoma FOCUS cells (Lavaissiere, *et al.*, 1990). One MCA reacted with an antigen that was highly expressed in hepatocellular carcinomas (Lavaissiere, *et al.*, 1996). Immunoscreening using this antibody and a lambda gt11 HepG2 library resulted in the isolation of a partial cDNA, which was subsequently used to isolate a larger clone.

A human adenocarcinoma cell line designated A549 was reported as having very high levels of HAAH activity (Lavaissiere, *et al.*, 1996). A mouse monoclonal antibody designated MCA 44-3A6 (U.S. Patent No. 4,816,402) was produced against the human adenocarcinoma cell line A549 (ATCC accession number CCL 185) (Radosevich, *et al.*, 1985). The antibody recognized a cell surface, non-glycosylated antigenic protein having an estimated apparent molecular weight of 40 kDa).

The antigen was expressed by A549 cells, and was found to be a good adenocarcinoma marker; that is, it was frequently expressed by cancers

which looked like adenocarcinomas when examined histologically (Radosevich, *et al.*, 1990a; Lee, *et al.*, 1985). MCA 44-3A6 is unique in that it is the first monoclonal antibody which has this binding specificity. The results from an International Workshop for Lung cancer confirmed other related published findings on MCA 44-3A6 (Stahel, 1994).

The antibody designated MCA 44-3A6 has clinical utility because it differentiates antigens associated with adenocarcinomas. The normal and fetal tissue distribution of the antigen is restricted to some glandular tissues (Radosevich, *et al.*, 1991). Detection can occur on formalin fixed-paraffin embedded tissue (Radosevich, *et al.*, 1985, 1988, 1990a, 1990b; Lee, *et al.*, 1985, 1986; Piehl, *et al.* 1988; Combs, *et al.*, 1988b, 1988c; Banner, *et al.*, 1985). The antibody has a restricted binding pattern within human pulmonary tumors (Lee, *et al.*, 1985; Banner, *et al.*, 1985; Radosevich, *et al.*, 1990a, 1990b).

In a study of over two hundred pulmonary cancers, MCA 44-3A6 was found to react with all of the adenocarcinomas tested, many of the large cell carcinomas, as well as with subsets of intermediate neuroendocrine small cell lung cancers, well-differentiated neuroendocrine small cell carcinomas, carcinoids, but not mesotheliomas. MCA 44-3A6 does not react with squamous cell carcinoma, bronchioloalveolar carcinoma, or small cell carcinoma (Lee, *et al.*, 1985). MCA 44-3A6 is useful in distinguishing adenocarcinomas that are metastatic to the pleura from mesothelioma (Lee, *et al.*, 1986). The antibody has selected reactivity among adenocarcinomas and in large cell carcinomas (Piehl, *et al.*, 1988; Radosevich, *et al.*, 1990b).

In a study of over 40 cases of lung cancer comparing cytological and histological findings, MCA 44-3A6 was found to be useful in cytological diagnosis and was consistent with the histological finding (Banner, *et al.*, 1985). Histology is the study of tissues (which are made of cells). Cytology is the study of cells which have been removed from the organizational context which is commonly referred to as tissue. Cells removed from tissues do not always behave the same as if they were in the tissue from which they were

derived. Fortunately, the antigen detected by MCA 44-3A6 expressed in adenocarcinoma cells in tissue behaves in the same ways as adenocarcinoma cells removed from tissues. This is a very diagnostically important characteristic. Similar correlations using cytologically prepared cell blocks of pulmonary carcinomas, as well as ACs presenting in body fluids from other sites throughout the body were demonstrated (Lee, *et al.*, 1985; Spagnolo, *et al.*, 1991; Combs, *et al.*, 1988c). Also, MCA 44-3A6 binds to adenocarcinomas from sites other than lung cancer. The expression of the antigen in primary and metastatic lesions was also reported (Combs, *et al.*, 1988a). The utility of the MCA antibody in differentiating cancer from benign lesions in human breast tissue was also noted (Duda, *et al.*, 1991).

The cellular localization of the antigen detected by MCA 44-3A6 was determined. By using live cell radioimmunoassays (a radioactive antibody test directed at determining binding of the antibody to live cells), immunofluorescence, and live cell fluorescence activated cell sorter (FACS) analysis, the antigen detected by MCA 44-3A6, was shown to be on the outside surface of the cell (Radosevich, *et al.*, 1985). Additional studies using immunogold-electron microscopy and FACS analysis have demonstrated that this antigen is non-modulated (that is not internalized by the cancer cell when bound by an antibody), is expressed on the extracellular surface of the plasma membrane, and is not cell cycle specific that is, the cell makes protein all the time it is going through the process of cell replication, and also when it is not dividing (Radosevich, *et al.*, 1991). The antigen is not found in the serum of normal or tumor bearing patients, and is not shed into the culture media by positive cell lines (that is, cancer cells are known to bleb off portions of their cell membranes and release them into the surrounding fluid.) (Radosevich, *et al.*, 1985). Recently 3 of 27 randomly tested adenocarcinoma patients were found to have naturally occurring antibodies to the antigen. In addition, radiolabeled MCA 44-3A6 was used to localize A549 tumors growing in nude mice. A doxorubicin immunoconjugate MCA 44-3A6 is selectively toxic in vitro (Sinkule, *et al.*, 1991).

Determination of the nucleotide and amino acid sequences of the antigen detected by MCA 44-3A6 would enhance the usefulness of this antigen in cancer diagnosis, treatment and prevention

BRIEF SUMMARY OF THE INVENTION

5 The antigen detected by the antibody MCA 44-3A6 as described in the Background is now designated as "Labyrinthin." A gene (designated *labyrinthin*; abbreviated *lab*) characterized by a unique nucleotide sequence that encodes the antigen detected by MCA 44-3A6 was isolated and characterized. (*lab* notation signifies the nucleic DNA/RNA forms; "Lab" notation refers to the protein which is encoded by the *lab* DNA/RNA).

10 The invention described herein used the antibody MCA 44-3A6 as a tool to clone the gene encoding Lab. In addition, an epitope (the necessary binding site for an antibody found on the antigen) for MCA 44-3A6 was identified on the Lab protein expressed by the clone to be PTGEPQ.² The epitope represents an important immunodominant sequence; that is, when injected into animals, the animals readily produce antibodies to this sequence.

15 An aspect of the invention is the use of *lab* DNA in the sense³ expression mode for: 1) the marking of human tumors by nucleotide probes; 20 2) the detection of DNA and mRNA expression of *lab* in cells and tissues; 3) the transformation of cells into a glandular-like cell type; 4) the production of Lab antigen *in vivo* for immunization; 5) the *ex vivo* expression of Lab for immunization to produce antibodies; and 6) production of Lab *in vitro*. Use of an antisense molecule, e.g. by production of a mRNA or DNA strand in the reverse orientation to a sense molecule, to suppress the growth of 25 *labyrinthin*-expressing (cancerous) cells is another aspect of the invention.

² Standard abbreviations for amino acids.

³ The normal transcription of a DNA sequence which proceeds from the 3' to the 5' end to produce a mRNA strand from the sense strand of DNA, the mRNA being complementary to the DNA.

An aspect of the invention is a vector comprising a DNA molecule with a nucleotide sequence encoding at least an epitope of the Lab antigen, and suitable regulatory sequence to allow expression in a host cell.

Another aspect of the invention is an amino acid sequence deduced from the protein coding region of the *lab* gene. Select regions of the sequence were found via immunological methods, to correspond and react to both naturally occurring (from cancer cells), chemically produced (synthetically produced peptides), and the expression of the cloned *lab* gene.

Another aspect of the invention is the use of the entire deduced amino acid sequence of Lab, peptides derived from Lab, or chemically produced (synthetic) Lab peptides, or any combination of these molecules, for use in the preparation of vaccines to prevent human cancers and/or to treat humans with cancer. For purposes of the present invention, "humans with cancer" are those persons who have the Lab antigen detected on their cells. These preparations may also be used to prevent patients from ever having these tumors prior to their first occurrence.

Monoclonal antibodies directed to the Lab protein, or antigen components or derivatives of Lab proteins, are useful for detection of Lab and for other purposes. Monoclonal antibodies which are made in species other than those which react with the Lab antigen can be modified by a number of molecular cloning such that they retain their binding with the Labyrinthin peptides, yet are not immunogenic in humans (Sastry, *et al.*, 1989; Sambrook, *et al.*, 1990). In brief, this is done by replacing the binding site sequence of a cloned human antibody gene, with the binding site sequence of the non-human monoclonal antibody of interest. These "humanized" MCAs are used as therapeutic and diagnostic reagents, *in vivo*, *ex vivo*, and *in vitro*.

The use of the Lab protein or antigenic peptides derived therefrom in diagnostic assays for cancer is a way to monitor patients for the presence and amount of antibody that they have in their blood or other body fluids or tissue. This detection is not limited to cancers of a class or classes previously defined, but is useful for cancer cells that have the Lab marker

antigen. The degree of seroconversion, as measured by techniques known to those of skill in the art [e.g. ELISA (Engrall and Perlmann, 1971)] may be used to monitor treatment effects.

Treatment with antisense molecules to *lab* or antibodies to Lab is an approach to treat patients who have Lab in, or on, their cancer cells.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 is the nucleic acid sequence of the *lab* gene. ^{SEQ ID NO: 1}

Figure 2 is the amino acid sequence for Lab, deduced from the *lab* gene. ^{SEQ ID NO: 2}

Figure 3 is an illustration of the *lab* gene and how it is related to the HAAH enzyme.

DETAILED DESCRIPTION OF THE INVENTION

Molecular Biology of *Labyrinthin*: To demonstrate that the MCA 44-3A6 epitope is encoded by a protein sequence, high molecular weight DNA from A549 was isolated. This DNA was co-precipitated (via calcium) with a plasmid (pSVneo), and used to transfect a mouse cell line designated B78H1 cells (Albino, *et al.*, 1985). This mouse cell line is negative for the expression of the epitope and was reported to have a high frequency of incorporation and expression of any human DNA sequences. If a given B78H1 cell was in a state to take up DNA, it would be expected to have taken up both human DNA and the plasmid DNA. The plasmid DNA makes the cell resistant to G418 (a normally toxic drug). Therefore, if a cell normally sensitive to G418 growth inhibitor grows in G418, it had to have taken up the plasmid, and may also have taken up one or more A549 DNA sequences. After G418 selection (a way of choosing only cells which have resistance to growth in G418 by the uptake/expression of the Neo gene on pSVneo plasmid, and therefore representing cells that were in a state to uptake other DNA at the same time), approximately 15 of 1×10^5 clones were detected using immunoselection with MCA 44-3A6. This finding is consistent

with a conclusion that human A549 cells have DNA that encodes Lab and possesses the regulatory sequences necessary for the expression of Lab.

Comparison of HAAH and *Labyrinthin*: Because the DNA sequence of *lab* was determined as an aspect of the present invention, HAAH and *lab* could be compared. HAAH and the *lab* nucleotide sequences have some internal fragment similarities, but are different on either side of the fragment, and are related to different products. This conclusion is based in part by the analysis and homology of the DNA sequences reported for these two genes. Specifically, the *lab* 5' region has no homology with HAAH. The protein coding region of *lab* has about a 99.6% homology with an internal segment of the proposed protein coding region for HAAH. The 3' region has no homology with the HAAH reported sequence. Virtually all of the other data comparing HAAH and *labyrinthin* are different, for example: 1) molecular weights of the proteins, 2) cellular localization, 3) chromosome localization, 4) histological presentation in normal tissues and tumors, 5) northern blot expression, 6) immunological findings.

Although the protein coding region of *lab* is identical to an internal region of the sequence reported for HAAH, the 5' untranslated region of HAAH is different, and part of the 5' translated protein coding region of HAAH is missing from that found in the *lab* clone. From both HAAH and *lab* clones, the deduced protein would be expected to be very acidic in nature, and therefore would run anomalously in SDS gels. As predicted, the Lab protein migrates anomalously in SDS gels. What was cloned and disclosed in the present invention migrates identically to the native protein found in several cell lines. Convincing evidence that the correct gene fragment encoding the antigen detected by MCA 44-3A6 has been cloned (mRNA) is that when the recombinant protein is made, that recombinant protein should act (in this case - have an apparent molecular weight) the same as independent biologically derived source of that protein. Lab provided from clones has the characteristics of Lab from cells.

The deduced amino acid sequence encoded by HAAH requires the use of an open reading frame which would produce a protein that is 85-90 kilodaltons, and does not take into account that there are several start codons and other shorter open reading frames. The deduced HAAH protein (biochemically) is glycosylated and the reported sequence has glycosylation sites (Korioth, *et al.*, 1994; Lavaissiere, *et al.*, 1996). To the contrary, Lab is not glycosylated, nor does it have predicted glycosylated sites.

The deduced HAAH amino acid sequence contains a region shared by the Lab amino acid sequence which is predicted to be very hydrophobic. Lab requires a strong detergents in order to be soluble; HAAH does not. The increased expression of HAAH (by enzyme activity measurements) in the same cell line (A549) which was used to clone and study lab extensively, suggests that both of these gene products may be important to the AC phenotype and that at least A549 cells make both functional HAAH and Lab. Successful transfections of the antisense to *lab* into A549 resulted in a marked decrease in expression of *lab* and in the growth rate of the cells. The expression of a sense *lab* construct in NIH-3T3 cells (normal mouse fibroblasts) resulted in a marked change in phenotype, a phenotype consistent with that of ACs. Therefore, *lab* expression is associated with conversion of normal cells to cancerous cells. Lab and HAAH have potential calcium binding domains in common.

cDNA Library Construction and Cloning: A cDNA lambda gt11 phage library was constructed using mRNA which was isolated from actively growing A549 cells (Sambrook, *et al.*, 1990). This oligo(dT)-primed cDNA was cloned into the Eco RI site using Eco RI linkers. The library has about 83% clear (containing an insert) plaques with a titer of 1.2×10^{10} /ml representing a minimum of 1.46×10^6 independent plaques which, by Polymerase Chain Reaction, have insert sizes ranging from 0.6 to 5 kilobases. Since Lab is a 40 kilodalton integral protein, (a protein which is embedded in the plasma membrane,) the theoretical full length mRNA encoding this protein, including a potential leader sequence would be about 1.1 kilobases. This library was

immunoscreened using the antibody MCA 44-3A6. Eight independently derived phage stocks (identical phage which are from the same plaque) were isolated. These have all been plaque purified by repeated cycles of immunoscreening/ isolation. Upon Eco RI digestion of these eight isolates, inserts of about 2 kb were seen. The largest insert was isolated (2A1A1) and the Eco RI fragment was cloned into the pGEM-3Z plasmid.

Sequencing and Sequence Analysis: The DNA fragment designated 2A1A1 was found to have an insert of 2442 base pairs in length (FIG. 1), containing a 5' untranslated region, a ribosome binding site, and a start codon which would be expected to encode a 255 amino acid protein (FIG. 2). The 3' untranslated region is remarkable in that it contains only four instability sequences; ATTTA (Xu, *et al.* 1997). In addition there are sequences found in the very 3' end of mRNA's which result in adenylation of the mRNA (Sambrook, *et al.*, 1990). The lab sequence contains both a sub-optimal (ATTAAA) and optimal (AATAAA) poly-adenylation site. These are sequences found in the very 3' end of mRNA's which result in adenylation of the mRNA. This finding provides molecular data which supports the cellular and biochemical data that has been outlined above. (The HAAH clone has a poly A signal, but the whole 3' region has not been sequenced.)

A calcium binding site motif is noted in the Lab amino and sequence (FIG. 2), however, it is out of the known required structural context to be a binding site. In this case, the calcium limiting sequence is there, but it is not in a protein sequence context that is known to make it work as a binding site! Homology was noted with *lab* and an EST clone (designated # 05501) which represented only a portion of the 3' untranslated region and independently confirmed this portion of the sequence. Some internal fragment homology is also noted with HAAH, but the 5' untranslated and part of the 5' translated region is different (58 amino acids), as well as a major portion of the 3' coding region is missing in *lab* (FIG. 3).

Genomic DNA Cloning and Analysis: Using a PCR fragment representing the protein coding region of *lab* as a probe, a genomic lambda

FIX II library made from the human pulmonary fibroblast cell line WI-38 was screened. Ten primary plaques were isolated out of approximately 1×10^6 screened plaques. Using seven of these as target DNA, Polymerase Chain Reaction conditions were established with primers for the protein coding region, producing a 765 base pair fragment, the expected protein coding region for *lab*. On Northern blots (a method used to qualitatively assess mRNA) *lab* only detects one band noted at 2.7 kilobases. The recombinant protein made from the *lab* clone, when tested on Western blots (a method used to qualitatively define proteins) using MCA 44-3A6, has the same relative mobility as the Lab protein when made by A549 cells.

Lab and HAAH genes give different results in the proteins they encode. HAAH consistently gives two bands on Northern blot analysis (2.6 and 4.3 kilobases) suggesting that the 2.6 kilobase band is due to alternative splicing, i.e. the cell cuts and splices the mRNA. Also, if *lab* and HAAH are the same gene, HAAH should be detected in all tissues and cancer cell lines in which Lab is found. However, Lab is not seen on Northern blots of cell lines EMT6 or QU-DB, nor is there immunoreactivity in these cells; indicating that Lab mRNA is not made, and that Lab protein is not produced in these cells. Lab protein is rarely expressed in normal cells, where both the HAAH mRNA and HAAH protein have been reported to be expressed in almost every tissue studied.

mRNA Analysis: Northern blot analysis of the DNA fragment from the A549 cell line using *lab* cDNA as a probe identified a single band of about 2.7 kilobases. This is expected based on the cDNA (2442 base pairs) and a poly-A tail of about 300 base pairs. Northern blot analysis of the mouse cell line, EMT6, and of the human large cell carcinoma cell line, QU-DB, confirm that no transcript for *lab* is produced by these cells. This is consistent with immunoassays which are negative for *lab* expression on these cells.

Antisense and Sense cDNA Expression: The plasmid (pBK-CMV) (Sambrook, *et al.*, 1990) may carry either the sense or antisense full length cDNA *lab* into A549 and NIH 3T3 cells. An antisense molecule can be, for

example, a complementary sequence to a sense molecule that hybridizes with the sense molecule, preventing its expression. Using the MTT assay (Siddique, *et al.*, 1992) to assess the growth rate of A549 cells expressing antisense to *lab*, a marked reduction in growth rate was noted. The antisense transfected A549 cells appear to have a greater degree of contact inhibition. A detectable amount of *Lab* is reduced in these antisense transfected cells. NIH-3T3 cells convert from a fibroblast-like cell type morphology (large, thin spindle shaped) to a large, adenocarcinoma appearing cells (very round, plump) when sense expression occurs.

Chromosome Localization: The chromosome localization for *lab*, using full length cDNA as a probe via *in situ* hybridization (Sambrook, *et al.* 1990) is tentatively on chromosome 2q12-14, with possibly some reactivity to chromosomes 4 and 8. Using the same probe (the full length cDNA sequence of *lab*) and FACS sorted chromosomes (Lebo, *et al.* 1985) staining was also noted on chromosome 2, with weak staining on 4 and none on 8. The use of genomic clones will be of particular value in resolving these data because higher stringency hybridization conditions than that allowable for the cDNA, can be used, thereby reducing background signals. This is yet another proof that the correct gene was cloned and that the results are not due to a method artifact. There may be mutations in the genomic DNA of tumors and for the present invention, DNA was cloned from tumor cells (A549). Therefore, a mutated gene could have been cloned. However, that is not the case because the genomic DNA from a normal cell (DNA) produced the same sequence as what cloned as described herein. Therefore, a normal gene was cloned from A549 cells. The weak signals on chromosomes 4 and 8 are consistent with a pseudogene or a related gene. For example, HAAH has been reported to be on chromosome 8q12 by *in situ* hybridization, so this result on chromosome 8 could reflect the HAAH and *lab* sequence homology.

Protein Molecular Characterization of *Labyrinthin*: Previous work using Western blot analysis (a qualitative assay to assess antigens) has shown that the *Lab* antigen is a 40 kilodalton (by relative mobility) protein detectable in

A549 cells (Radosevich, *et al.*, 1985). The epitope does not appear to be modulated or blocked by lectins, and is selectively expressed on the cell surface, primarily localized to the plasma membrane. (Radosevich, *et al.*, 1985, 1991). Lab is sensitive to proteases, but not lipid or carbohydrate altering reactions (Radosevich, *et al.*, 1985). The biochemical properties of Lab are consistent with Lab being an integral membrane protein.

Having a deduced amino acid sequence from the *lab* gene of the present invention, allows further characterization of the Lab protein. Extensive computer analysis of Lab has identified a eukaryotic leader-like sequence and theoretical cleavage site, 3 myristylation sequence sites, a weak membrane anchoring domain (MAD I), and a strong membrane anchoring domain (MAD II) (FIG. 2). [(In the HAAH sequence, there are 58 (theoretical) amino acids followed by a sequence homology in the Lab protein coding sequence, and an additional 445 amino acid 3' to the *lab* sequence.]

When Lab is expressed as a fusion protein in a bacterial GST fusion expression system (pGEMEX-2T) (Amersham Pharmacia Biotech, Inc., Piscataway, New Jersey, 08854, USA), and subjected to Western blot analysis using the antibody MCA 44-3A6, the resulting blots demonstrate that the expressed cleaved fusion protein has the same relative mobility as the protein detected in A549 cells. The deduced molecular weight for Lab is 28.8 kilodaltons and on Western blots it has a relative mobility identical to the form expressed by A549 cells (apparent relative mobility = 40 kilodaltons). The 55 glutamic and 27 aspartic acid residues (82 residues combined) are almost uniformly distributed throughout the protein (255 amino acids total; 228 no leader sequence), except for the leader sequence and the strongest membrane anchoring domain (MAD II). These data suggest that Lab migrates anomalously in SDS gels. Cell lines other than A549 (e.g. adenocarcinomas DU-145, ATCC # HTB-81; ZR-75-1, ATCC # CRL-1504, and so forth) have an antigen detected with the same molecular weight antigen as Lab. Neither a 85-90 kilodalton molecular weight species, nor a

52 and 56 kilodalton molecular weight species is noted when probing Western blots for Lab.

Epitope Mapping Using the Antibody MCA 44-3A6 and Vaccine

Feasibility of Lab: Using Polymerase Chain Reaction and the GST fusion

5 protein system, subclones of the protein coding region were made, and epitopes mapped the binding of MCA 44-3A6 to six amino acids (PTGEPQ) representing amino acids #117-122 of Lab ("P" peptide). In order to determine this epitope, the entire coding region was divided into regions, Polymerase Chain Reaction primers were designed to amplify each region, and the subsequent expression of Polymerase Chain Reaction products were
10 cloned and tested by Western blot analysis using the antibody MCA 44-3A6.

The DNA fragment representing the positive Western blot result was then further subdivided. Polymerase Chain Reaction products were generated and cloned, expressed, and tested via Western blot. Constructs
15 were made in this way both from the 5' end and the 3' end and the intervals of the number of amino acids were reduced upon each round. This resulted in the last round representing a one amino acid difference from the previous round (in both directions), such that one could deduce the exact binding site of the MCA 44-3A6. This demonstrates that at least these six amino acids
20 are exposed to the external cell surface. To further prove the point, the DNA encoding only these six amino acids have been cloned and the fusion protein is positive by Western blot analysis. Synthetically prepared "P" peptide can be specifically detected by MCA 44-3A6, and the synthetic peptide was immunogenic in 5 of 5 mice tested. Computer analysis/ modelling also
25 predicted that this epitope would be very immunogenic using computer assisted analysis (GCG programs) (Genetics Computer Group, Madison, WI 53703).

Vaccine Preparation: A vaccine is a preparation of antigen(s), which
30 when given to a host, results in the host producing antibodies against the antigen(s). The host response results in the host being immune to the disease to which the vaccine was directed. Vaccine treatment therefore,

prevents the clinical presentation of a disease, without the host being exposed to the disease causing agents. Lab has all the characteristics of a preferred cancer vaccine. The *lab* gene is frequently expressed by tumors which look like adenocarcinomas, is expressed on the outside of the cells, is expressed by all of the cells within a given cancer, is expressed at all times by these cancer cells, and is infrequently expressed by normal cells. Lab protein (peptides) can be produced by any number of methods using molecular cloning techniques, and can be produced in large quantities, thus making it a practical antigen to use as a vaccine. After the Lab protein has been purified so that it is suitable for injection into humans, it is administered to individuals intradermally, subcutaneously, or by other routes, so as to challenge the immune system to produce antibodies against this protein (peptides).

The use of molecular modeling and computer assisted analysis GCG programs (Genetics Crystal Group, Madison, WI 53703) allows the identification of small portions of a molecule, slightly larger than an epitope (six to seven amino acids for proteins), which are expected to be on the surface of a protein molecule. In addition, the degree of hydrophobicity or hydrophilicity of a given sequence, and how immunogenic the sequence would be in animals, can be determined (Genetics Crystal Group, Madison, WI 53703). After defining which sequences meet these criteria, the peptides are synthetically made, or produced by a number of standard methods. One or more of these peptides can then be formulated to be used as a vaccine, and administered to the host as outlined above, as a vaccine.

A vaccine comprising a molecule having an amino acid sequence selected from the group of sequences encoded by the cDNA of FIG. 1, ^{SEQ ID NO: 6} sequences encoded by the cDNA of FIG. 2, the peptides APPEDNPVED₁, ^{SEQ ID NO: 7} EEQQEVPPDT₁, ^{SEQ ID NO: 8} DGPTGEPQQE₁, ^{SEQ ID NO: 9} and QENPDSSEPV₁ and any fragment, or combinations thereof.

A given vaccine may be administered once to a host, or may be administered many times. In order for some patients to recognize a given

vaccine, an adjuvant may also need to be administered with the peptides. Adjuvants are nonspecific immune stimulators which heighten the immune readiness and aid in the conversion of the host from not having detectable serum antibodies to having very high titer serum antibodies. It is this high level (titer) of antibodies, which effectively protects the host from the diseases or conditions to which the antibodies are directed.

Functional Studies: Studies directed at understanding the cellular function(s) of Lab are extensions of initially reported, cell localization/ characterization studies (Siddique, *et al.* 1992). Changes in levels of Lab in response to extracellular exposure to various cations (Ca^{++} , Mg^{++} , Cu^{++} , and Fe^{++}) were undertaken. Lab expression in A549 cells was only modulated by Ca^{++} . Using the highly specific fluorescent Fura-2/AM Ca^{++} method of measuring cytosolic Ca^{++} , (Molecular Probes Inc., Eugene, OR 97402) it was demonstrated that: 1) the internal Ca^{++} concentration is higher in A549 cells than in QU-DB cells, and 2) that the A549 cell line responds to various external Ca^{++} levels (Siddique, *et al.*, 1992). Since pH can modulate intracellular free Ca^{++} levels, external pH manipulations should result in changes in the expression levels of Lab. Extracellular pH changes (in the presence of normal Ca^{++} concentrations) result in 1) a parallel change in intracellular pH as measured by SNARF-1 AM/FACS, (Molecular Probes Inc., Eugene, OR 97402) 2) transcript levels increase for Lab (when compared to GAPDH expression via Northern blot), and that 3) Lab protein also increases (using Western/Slot blot analysis). The intracellular changes in pH (due to external changes) for A549 cells are identical to those reported for normal cells. The increased expression of *lab* is also not due to cell death (as measured by MTT assays) (Siddique, *et al.*, 1992). In addition, incubation of recombinant Lab at various pH solutions does not alter immunoreactivity. Preliminary data suggests that when these experiments are conducted on A549 cells grown in reduced Ca^{++} , the induced expression of *lab* is blunted.

Methods of Diagnosing Cancer Cells in a Sample of Cells: Biological samples from a subject are used to determine whether cancer cells are

present in the subject. Examples of suitable samples include blood and biopsy material. One method of diagnosis is to expose DNA from cells in the sample to a labeled probe that is capable of hybridizing to the *lab* gene, or a fragment thereof, under stringent conditions, e.g. 6x ssc; 0.05x blotto; 50% formamide; 42°C (Sambrook, *et al.*, 1990). Of course, the hybridizing conditions are altered to achieve optimum sensitivity and specificity depending on the nature of the biological sample, type of cancer, method of probe preparation, and method of tissue preparation.

After contacting the sample with the probe, the next step is determining whether the probe has hybridized with nucleotide sequences of the DNA from the sample, from which the presence of the *lab* gene is inferred, said presence being diagnostic of cancer.

Another diagnostic method is to obtain monoclonal antibodies preferably labeled, either antibodies already existing, or new ones directed to the antigenic peptides that are aspects of the present invention, and contact a sample with these to detect the Lab antigen. These monoclonal antibodies are useful in the development of very specific assays for the detection of Lab antigen, and allow the tests to be carried out in many different formats; resulting in a broader application in science and medicine.

The current invention is useful in that it describes a new gene which is expressed on the surface of tumors, which was not previously reported. This gene is not tissue specific, and therefore will allow the detection of tumors regardless of the organ in which they arise. Likewise, the use of this gene to produce a vaccine for these tumors, will have a very broad application.

Diagnostic tests will also have this broad tissue use, making the detection of Lab//*lab* a "pan-marker" for cancer, in particular for what have been designated previously, adenocarcinomas.

DOCUMENTS CITED

Albino, AP, Graf, LH, Kontor, RRS, *et al.* DNA-mediated transfer of human melanoma cell surface glycoprotein gp130: Identification of transfectants by erythrocyte rosetting. *Mol. Cell. Biol.* 5:692-697, 1985.

5 Banner BF, Gould VE, Radosevich JA, *et al.* Application of monoclonal antibody 44-3A6 in the cytodiagnosis and classification of pulmonary carcinomas. *Diag Cytopathol.* 1:300-307, 1985.

Brown, DT and Moore, M. Monoclonal antibodies against two human lung carcinoma cell link. *Br. J. Can.* 46:794-801, 1980.

10 Combs SG, Hidvegi DF, Ma Y, *et al.* Pleomorphic Carcinoma of the Pancreas: A rare case report of combined histological features of pleomorphic adenocarcinoma and giant cell tumor of the pancreas. *Diag. Cytopathol.* 4:316-322, 1988a.

15 Combs SG, Radosevich JA, Ma Y, *et al.* Expression of the Antigenic Determinant Recognized by the Monoclonal Antibody 44-3A6 on Select Human Adenocarcinomas and Normal Human Tissues. *Tumor Biol.* 9:116-122, 1988b.

20 Combs SG, Radosevich JA, and ST Rosen. Cytological expression of the adenocarcinoma antigen marker in human body fluids. *Tumor Biol.* 9:116-122, 1988c.

Duda RB, August CZ, Radosevich JA and ST Rosen. Monoclonal Antibody 44-3A6 as a Marker For Differentiation of Breast Cancer. *Tumor Biol.* 12:254-260, 1992.

25 Engvall, E and Perlmann, P. Enzyme linked immunosorbent assay (ELISA): Quantitative assay of IgG. *Immunochemistry.* 8:87-874, 1971.

Gronke RS, VanDusen WJ, Garsky VM, Jacobs JW, Sardana MK, Stern AM, and PA Friedman. Aspartyl beta hydroxylase: In vitro hydroxylation of a synthetic peptide based on the structure of the first growth factor-like domain of human factor IX. *PNAS.* 86:3609-3613, 1989.

Gronke RS, Welsch DJ, VanDusen WJ, Garsky VM, Sardana MK, Stern AM, and PA Friedman. Partial purification and characterization of bovine liver aspartyl beta hydroxylase. J. Biol. Chem. 265:8558-8565, 1990.

5 Jia S, VanDusen WJ, Diehl RE, Kohl NE, Dixon RAF, Elliston KO, Stern AM, and PA Friedman. cDNA cloning and expression of bovine aspartyl (asparagineinyl) beta-hydroxylase. J. Biol. Chem. 267:14322-14327, 1992.

10 Jia S, McGinns K, VanDusen WJ, Burke CJ, Kuo A, Griffin PR, Sardana MK, Elliston KO, Stern AM, and PA Friedman. A fully active catalytic domain of bovine aspartyl (asparagineinyl) beta-hydroxylase expressed in Escherichia coli: Characterization and evidence for the identification of an active-site region in vertebrate alpha-ketoglutarate-dependent dioxygenases. PNAS 91:7227-7231, 1994.

15 Korioto F, Gieffers C, and J Frey. Cloning and characterization of the human gene encoding aspartyl beta-hydroxylase. Gene 150:395-399, 1994.

Landis, S.H., Murray, T., Bolden S., and P.A. Wingo. Cancer Statistics, 1998., CA 44:6-9.

20 Lavaissiere L, Jia S, Nishiyama M, de la Monte S, Stren AM, Wands JR, and PA Friedman. Overexpression of human aspartyl (asparagineinyl) beta-hydroxylase in hepatocellular carcinoma and cholangiocarcinoma. J. Clin. Invest. 98:1313-1323, 1996.

25 Lebo, RV, Tolan, DR, Bruce, BD, Cheng, MC, and Kan, YW. Spot blot analysis of sorted chromosomes assigns a fructose intolerance gene locus to chromosome 9. Cytometry. 6:476-483, 1985.

Lee I, Radosevich, JA, Rosen, ST, *et al.* Immunohistochemistry of lung carcinomas using monoclonal antibody 44-3A6. Can. Res. 45:5813-5817, 1985.

30 Lee I, Radosevich JA, Chejfec G, *et al.* Malignant Mesotheliomas: Improved Differential Diagnosis From Lung Adenocarcinomas Using

Monoclonal Antibodies 44-3A6 and 624A12. *Amer. J. Path.* 123:497-507, 1986.

Livingston, PO, Wong, GYC, Adluri, S, Tao, Y, Padevan, M, Parente, R, Hanlon, C, Calves, MJ, Helling, F, Ritter, G, Oettgen, HF, and Old, LJ.
5 Improved survival in AJCC stage III melanoma patients with GM2 antibodies: A randomized trial of adjuvant vaccination with GM2 ganglioside. *J. Clin. Oncol.*, 12:1036-1044, 1994.

Piehl MR, Gould VE, Radosevich JA, *et al.* Immunohistochemical Identification of Exocrine and Neuroendocrine Subsets of Large Cell Lung
10 Carcinomas. *Path. Res. and Prac.* 183:675-682, 1988.

Radosevich JA, Ma Y, Lee I, *et al.* Monoclonal antibody 44-3A6 as a probe for a novel antigen found human lung carcinomas with glandular differentiation. *Can. Res* 45:5805-5812, 1985.

Radosevich JA, Lee I, Gould VE, and ST Rosen. Monoclonal antibody
15 assays for lung cancer. In: *In vitro diagnosis of human tumors using monoclonal antibodies.* Kupchik HZ and N Rose (Eds.) Marcel Dekker p101-119, 1988.

Radosevich JA, Combs SG, and ST Rosen. Immunohistochemical analysis of lung cancer differentiation markers. In: *Lung Cancer Differentiation. Lung Biology in Health and disease.* L'Enfant C, Bernal S, and Baylin S. (Eds.). Marcel Dekker, 1990a.
20

Radosevich JA, Noguchi M, Rosen ST, Y. Shimosato. Immunocytochemical analysis of human adenocarcinomas and bronchioloalveolar carcinomas of the lung using the monoclonal antibody
25 44-3A6. *Tumor Biology.* 11:181-188, 1990b.

Radosevich JA, Combs SG, and ST Rosen. Expression of MCA 44-3A6 in human fetal development. *Tumor Biology* 12:321-329, 1991.

Radosevich JA, Siddique FS, Rosen ST, and WJ Kabat. Cell Cycle and EM Evaluation of the Adenocarcinoma Antigen Recognized by the
30 Monoclonal Antibody 44-3A6. *Br. J. Can.* 63:86-87, 1991.

Rosen, ST, Mulshine, JL, Cuttitta, F, and Abrams, PG. Biology of Lung Cancer. Marcel Dekker, Inc. New York, NY, Vol. 37, 1988.

Sambrook J, Fritsch EF, and T Maniatis. Molecular cloning: a laboratory manual. 2nd Ed. Cold Spring Harbor Lab. Press., 1990.

5 Sastry, L., Alting-Mees, M, Huse, WD, Short, JM, Hay, BN, Janda, KD, Benkovis, SJ, and Lerner. Cloning of the immunological repertoire in *Escherichia coli* for generation of monoclonal catalytic antibodies: Construction of a heavy chain variable region-specific cDNA library. PNAS. 86:5728-5732, 1989.

10 Siddique FS, Iqbal Z, and JA Radosevich. Changes in the expression of the tumor-associated Antigen recognized by monoclonal antibody 44-3A6 in A549 cells due to calcium. Tumor Biol. 13:142-151, 1992.

15 Sinkule J, Rosen ST, and JA Radosevich. MCA 44-3A6 Doxorubicin (Adriamycin) Immunoconjugates: Comparative In Vitro Anti-Tumor Efficacy of Different Conjugation Methods. Tumor Biol. 12:198-206, 1991.

20 Spagnolo DV, Witaker D, Carrello S, *et al.* The use of monoclonal antibody 44-3A6 in cell blocks in the diagnosis of lung carcinoma, carcinomas metastatic to lung and pleura, and pleural malignant mesothelioma. Am. J. Clin. Path. 95:322-329, 1991

25 Stahel, RA (Chairman). Third International IASLC Workshop on Lung Tumor and Differentiation Antigens. Inter. J. Cancer Suppl 8:6-26, 1994.

 Wang Q, VanDusen WJ, Petroski CJ, Garsky VM, Stern AM, and PA Freidman. Bovine liver aspartyl beta-hydroxylase. J. Biol. Chem. 266:14004-14010, 1991.

30 Xu, N., Chen, C-Y A, Shyu, A-B. Modulation of the fate of cytoplasmic mRNA by AU-rich elements: Key sequence features controlling mRNA deadenylation and decay. Mol. Cell. Biology. 17:4611-4621, 1997.